Critique of the PQRI/BUWG Documents

R. Gary Hollenbeck
Associate Dean and Professor of Pharmaceutical Sciences
University of Maryland School of Pharmacy

Demonstrating Adequacy of Mix

Sometimes in the midst of a controversy the obvious gets lost: a uniform distribution of the active pharmaceutical ingredient(s) throughout a blend is a good thing. There is no disagreement that blend uniformity is important, and pharmaceutical scientists recognize that it is unlikely that content uniformity of the dosage form will be achieved when the blend is not adequately mixed. The manufacturers and regulators agree on the importance of blend uniformity; it is the testing aspects of this issue that challenge us.

With regard to in-process blend testing, using some sort of interventionist sampling device, the key question has been: Does blend uniformity data so acquired consistently correlate with dosage form uniformity data? The data-mining effort conducted by the PQRI Blend Uniformity Working Group has clearly shown that the answer is "not always." Given this outcome, it is inappropriate for FDA to require—directly or indirectly—use of a frequently unreliable test.

With regard to post-process testing of dosage units, who really believes that a test on 20 tablets is sufficient to characterize the content uniformity of a lot with hundreds of thousands of tablets or capsules?

The proposal to use stratified sampling of dosage units as an alternative to routine blend sample analysis to demonstrate adequacy of mix for powder blends is a science-based, next generation approach that makes sense:

- It maintains the ability to use the existing system when it has been or can be shown to be appropriate
- It places the responsibility to justify and support an alternative approach on the industry during process development
- It results in increased confidence that there is content uniformity of dosage units in large scale production

The dosator on an encapsulator or the die cavity on a tablet machine are the quintessential blend-sampling devices for the encapsulation and tablet compression unit operations. Strong arguments can be made that the resulting capsule or tablet is the most appropriate sample for assessment purposes. The stratified sampling proposal of 20 uniformly spaced points throughout a filling or compression process offers the added dimension of characterizing time-dependent processes such as stratification during a run.

A positive response from FDA to this science-based analysis and initiative would insure the quality of manufactured products and validate the benefits to society of a collaborative PQRI.

Critique of the PQRI/BUWG Documents

K.J. Leiper, C.Chem., MRSC
Benson Associates, Caythorpe NG323DJ

The PQRI Blend Uniformity Working Group's (BUWG) papers are two of *the* most eagerly awaited publications for some time, not just to evaluate the part these studies might ultimately play in bringing the revision and acceptance of the FDA's Guidance on Blend Uniformity Analysis to a successful conclusion but perhaps more importantly to provide a measure of the contribution that could reasonably be expected of PQRI projects in future.

While the authors claim, "The resulting recommendation addresses both industry and FDA concerns without compromising product quality or increasing regulatory burden," the acid test will be whether it will contribute to a cGMP/ guidance which significantly reduces in non-compliance observations from the regulatory perspective and deliver a consistent approach in the ANDA assessment as well as manufacturing efficiency improvements for industry. Can these recommendations satisfy both parties' aspirations?

One can of course relatively easily try to assess how regulators and industry might view these recommendations by challenging them with questions such as:

- Is chemically based univariate determination on the dosage form the best way of assessing the performance of a dynamic multivariate system?
- Are data on 148 batches from 8 manufacturers covering only one of the three areas under investigation really representative of industry performance?
- How can we effectively control the physical variability of input material on process performance across the product life cycle—continue to ignore it?
- Which industry today is turning toward increased product testing to manage its process risks and manufacturing efficiencies?
- With moves toward higher product testing regimes how will increasing numbers of outliers be handled—using OOS investigations?
- Is increased post mortem product analysis or, in the worst case, a market recall still a realistic option?
- How relevant are compendial requirements and specifications to control product quality?
- Is the compendial requirement of testing 30 tablets for uniformity and 6 for dissolution scientifically justifiable or relevant?
- Has normal distribution not been proven—not assumed?
- What are the real statistical sample requirements to attempt to exercise product-based control on a batch of a million tablets?
- Are there any other more viable alternatives, etc.?

The fact is that while these investigations have been ongoing the goalposts have moved. The issues, which need to be addressed, are not just sampling related. Regulators and industry now recognize that the only real resolution of blending issues lies in improved process understanding. Overemphasis on the symptoms will only serve to further delay progress toward real resolution of the cause—understanding and controlling the blending process—or even more radically by defining viable novel processing alternatives.

The industry certainly needs initiatives in the sphere of Product Quality Research spanning industry and regulatory needs, but there is also a requirement to be aware of the broader implications. Perhaps against such criteria the BUWG recommendations may now be seen as too focussed and tactical.

Critique of the PQRI/BUWG Documents

Chris Moreton, PhD.
Consultant to the Pharmaceutical Industry and a Member of the PQRI Drug Product Technical Committee.
Toronto, OH.

The Blend Uniformity Working Group (BUWG) reports were submitted to the FDA in February 2002. Recently (May 2002) the Agency announced the withdrawal of the 1999 draft Guidance on blend uniformity testing. The Agency appears to have accepted that blend uniformity testing can be flawed. The aim of the data-mining analysis was to verify that the assumptions made in the Mote Carlo simulations were valid, and to determine if blend uniformity was a useful predictor of content uniformity. The data analysis showed that for most data points the assumption of normality was valid. Where normality was not found, the associated errors were on the side of caution. The report concludes correctly that the assumption of normality in the Monte Carlo simulations was valid and would not have adversely biased the outcome. The data analysis also shows that blend uniformity testing did not correlate with content uniformity either for the whole dataset or for subsets. The report properly concludes that blend uniformity testing is not a reliable predictor of content uniformity, and that there may be better ways to demonstrate adequacy of mixing. Data submitted were for tablet products; no data on hard shell capsule or powder products were submitted. This could hinder acceptance of the blanket proposal for solid dosage forms. However, the manufacture of blends for capsule filling use the same type and size of equipment and processing as tablet blends.

The proposed alternative to blend uniformity testing in routine manufacture is stratified sampling and subsequent testing of the output stream during the compaction or filling operation. However, this is predicted on having both blend uniformity and content uniformity data from stratified sampling at the development and validation stages, and validated process. The stratified sampling proposals distinguish between routine and exhibit/validation batches. Sampling and testing is more extensive for between routine and exhibit/validation batches, as might be expected. Stratified sampling targets those parts of an operation that could be at an increased risk of failure, together with monitoring of the output over the duration of manufacture. Sampling at regular intervals throughout the manufacturing operation will demonstrate compliance more reliably than blend uniformity testing. The wording of 21 CFR §211.110 (a)(3) requires that drug product manufacturers establish written procedures including in process controls 'where appropriate' and including 'Adequacy of mixing to assure uniformity and homogeneity'. With the withdrawal of the draft Guidance there is nothing specifically mandating blend uniformity testing. Thus, the BUWG proposals would comply with the spirit and letter of 21 CFR §211.10 (a)(3). The BUWG proposals are a serious attempt to provide a solution to a current problem. The proposals are an acceptable basis for a new guidance that will eventually be issued. The testing is directed to those parts of the operation where it will be most meaningful. This will ultimately benefit the patient, industry and the FDA by giving added assurance on the uniformity of marketed drug products.

Critique of the PQRI/BUWG Documents

Dave Rudd

It is encouraging that the PQRI group is providing recommendations, which are clearly described as 'an alternative approach'. We should not rule out (where applicable) the conventional approach nor, more importantly, any revised approach based on real-time monitoring of powder blend homogeneity (using near infra-red or solid-state fluorescence spectroscopy, for example). In general, the proposals make good scientific sense, although there are points of details (see 'specific comments', below), which raise some questions. The wider implications of these papers need to be carefully understood—for example, the influence of (discredited) powder blend sampling on the ability to carry out process validation; the impact of the basis of the proposals on 'Quality by Design' and Parametric Release concepts; the increase in workload for routine QC and/or Process Analytical laboratories, etc.

Specific comments on "The Use of Stratified Sampling…" paper:

- III. Background (final bullet point 3) - "It accounts for segregation after blending"

Whether such segregation works in favor of or against content uniformity, it is unacceptable to develop manufacturing processes which have uncontrolled elements such as this. If segregation is occurring, either it needs to be eliminated, or demonstrated to be consistent (which is unlikely). The admission that registered manufacturing processes feature such lack of control is not acceptable.

- IV. Process development (midway through first paragraph)
  - "Appropriate blend sampling techniques and procedures should be developed..."

Is this not the very issue which the PQRI is trying to overcome with these two papers? If blend sampling can be achieved during process development and validation, why can't the same procedure be used during routine batch manufacture?

Specific comments on "Results of Statistical Analysis…" paper:

- Notes for in-process blend testing results

A series of sample size ranges are described, but there is no further mention of these in the 'Results and Discussion' section. What actual sample sizes did the eight pharmaceutical companies use, as this would have a major bearing on the quality and nature of the data supplied?

- Normality of between-location means

The test for normality is not a valid one as the Wilk-Shapiro test was conducted on data means taken from a series of sample locations. The Central Limit Theorem says that "the sampling distribution of the mean tends towards the normal distribution as *n* increases, *even if the original population is not normally distributed.*" (Miller and Miller, Ellis Horwood Limited, 1984 - ISBN 0-85312-662-3). Thus, the fact that means are normally distributed does not necessarily confirm that the original data population is also normally distributed.

- Relationship between blend and dosage form RSD

Which blend sampling methods were used? What sample sizes were taken? Can we have confidence in any of these results without knowing the experimental details?

If we are to assume confidence in the data, rather than use the results to support the PQRI proposals, should we not use the results to support the status quo, as arguments around the difficulty of powder blend sampling and subsequent segregation strangely no longer seem to apply here?

- Predictive relationship of blend criteria for dosage form PDA criteria – "The dataset analyzed …demonstrate that there is no predictive value in any of the above blend criteria in regard to meeting the dosage form criterion"

How does PQRI feel about this statement when compared to that of the previous document (under IV. Process development) which states that "In general, content uniformity of the final dosage form is dependent on the homogeneity of the powder mixture in the blender"?

Intuitively, there must be a relationship between powder blend homogeneity and the content uniformity of the final dosage form. If blend criteria fail to provide a predictive relationship, this is telling us something about the shortcomings of the blend criteria and/or the propensity for segregation during post-blend processing. This is, therefore, where attention needs to be directed.

- Use of dosage unit uniformity data to demonstrate blend adequacy of mix (end of first paragraph)
  – "It is possible that this unreliability may be due to difficulties encountered when sampling the blend, which can bias the results "

This leads to the constructive suggestion that powder blend sampling issues are probably best overcome by the avoidance of sampling; hence, emphasis on non-sampling measurement techniques such as near infrared spectroscopy (NIR) or solid-state fluorescence spectroscopy. Incidentally, NIR is described as an 'emerging' technique. The application to on-line powder blend homogeneity testing is now well-established (there are several publications), such that this description may be a little out-dated. It is a mature technique for this type of application.

Critique of the PQRI/BUWG Documents

Ken Morris, PhD.,
Associate Professor, Department of Industrial and Physical Pharmacy
Purdue University,
West Lafayette, IN 47907-1336
morris@pharmacy.purdue.edu

The "Results of Statistical Analysis of Blend and Dosage Unit Content Uniformity Data Obtained from the Product Quality Research Institute Blend Uniformity Working Group Data-Mining Effort" represents a great effort to dimensionalize the issue of blend vs. dosage form uniformity sampling for content uniformity. The conclusions are based on the analyses of an unusually broad and unique data set. The reported trends follow common experience (i.e. blend uniformity differs from dosage form uniformity in high %CV blends, and analyzing the dosage form is the most meaningful measure of uniformity) and as such, circumstantially support the data. There are several concerns that vary in possible importance depending upon whether or not the team actually addressed the issues but just did not include the findings/details in the paper. These include:

"*Results and Discussion, Validation of Assumptions Used for Computer Simulations, Normality of Between Location Means*," third paragraph – The assumption of normal distributions is said to give rise to conservative results "…rejection rate estimates that are slightly smaller (more conservative) than criteria rejection rates based on actual data…" While statistically this may be more conservative, it sounds like the method is rejecting fewer samples than the compendial method. This makes it *less* conservative from the FDA point of view if we're reading this correctly. For example, the conservative issue is not well supported by some of the percentages given in Table 2.

"*Results and Discussion, Validation of Assumptions Used for Computer Simulations*," first paragraph – The paragraph states, "…extensive use was made of Mote Carlo simulation." However, no details or citation of the algorithm used are given. We realize that this is not a paper on Monte Carlo (MC) methods, still it is not clear how or even why it is used. We feel the work would greatly benefit from a relatively non-mathematical explanation of the steps (maybe even a diagram?). At the very least the identity and/or description of the algorithm is needed. Without this, the connection between the "proof of normal distribution" in the form of Shapiro-Wilk method applied to the data and the *actual* method of powering the study (i.e. the MC method) is lost. This could be a significant stumbling block for acceptance. Also, without a more complete MC discussion, it will be difficult to convince anyone of the validity of using the same sampling regime and level irrespective of batch size.

Data – It appears that all dosage forms were pooled. It would be useful to treat the DC, wet granulation, and dry granulation, data separately as well as pooled. This could serve

to support the approach by agreement with our "common experience" and to illustrate the ability of the method to discern trend differences.

Tables – Many of the tables are under-described in the text. As mentioned above, on Pg, 6 Table 2 appears to contradict the normality assumption, where some simple explanation may explain this easily. This is true for other tables, as well.

Methods – There is no mention of the analytical techniques used to analyze the blends or dosage forms. I am assuming they used either a standard UV technique or some provided by the companies. In either case this should be included.

There is also no detail on how the blends were sampled. Again, this should be included for completeness.

General – This document will have to be formatted as an article of course before publication. Even as an internal working document it should be reorganized a little more like the final publication.